

# PARTIALLY OBSERVABLE RISK-SENSITIVE STOPPING PROBLEMS IN DISCRETE TIME

NICOLE BÄUERLE\* AND ULRICH RIEDER†

**ABSTRACT.** In this paper we consider stopping problems with partial observation under a general risk-sensitive optimization criterion for problems with finite and infinite time horizon. Our aim is to maximize the certainty equivalent of the stopping reward. We develop a general theory and discuss the Bayesian risk-sensitive house selling problem as a special example. In particular we are able to study the influence of the attitude towards risk of the decision maker on the optimal stopping rule.

**KEY WORDS:** optimal stopping, certainty equivalent, exponential utility, updating operator, value iteration

**AMS SUBJECT CLASSIFICATIONS:** Primary 62L15, Secondary 90C40

## 1. INTRODUCTION

In this paper we consider stopping problems with partial observation under a general risk-sensitive optimization criterion for problems with finite and infinite time horizon. More precisely our aim is to maximize the certainty equivalent of the stopping reward over the time horizon. In case of an infinite time horizon we assume that we have strictly negative cost when we do not stop. The certainty equivalent of a random variable is defined by  $U^{-1}(\mathbb{E}U(X))$  where  $U$  is an increasing concave function. If  $U(x) = x$  we obtain as a special case the classical risk-neutral decision maker. The case  $U(x) = \frac{1}{\gamma}e^{\gamma x}$  is often referred to as 'risk-sensitive', however the risk-sensitivity is here only expressed in a special way through the risk-sensitivity parameter  $\gamma \neq 0$ . More general, the certainty equivalent may be written (assuming enough regularity of  $U$ ) as

$$U^{-1}\left(\mathbb{E}[U(X)]\right) \approx \mathbb{E}X - \frac{1}{2}l_U(\mathbb{E}X)\text{Var}[X] \quad (1.1)$$

where

$$l_U(x) = -\frac{U''(x)}{U'(x)} \quad (1.2)$$

is the *Arrow-Pratt* function of absolute risk aversion. In case of an exponential utility, this absolute risk aversion is constant (for a discussion see [5]).

In addition we suppose that the process is partially observable. More precisely we suppose that we have a jointly Markovian process  $(X_n, Y_n)$  where only the first component is observable. Also the stopping reward depends only on the first component. This class of models includes in particular Bayesian models where the unobservable process reduces to an unknown parameter. Important applications are sequential probability ratio tests. A particular example can e.g. be found in [22]. The theory of risk-neutral stopping problems can e.g. be found in [8] or [18].

Stopping problems with general utility functions are rarely treated in the literature. [17] considers the classical house selling problem with general utility but with complete observation. In a separate section we show that most of the results in [17] also extend to the Bayesian case. [15] considers stopping problems with denumerable state space and arbitrary utility function. The authors there discuss the so-called *monotone case* and give conditions for the optimality of *one-step-look-ahead rules*. Of course the stopping problems we treat here can be seen as special partially observable risk-sensitive Markov Decision Processes. The general theory for these type of problems has been developed in [2]. The special case of an exponential utility combined with

partial observation has been treated in [6], [11], [14], [19]. The first treatment of the exponential certainty equivalent as an optimization criterion can be traced back to [13]. See also [20] and [21] for other risk sensitive criteria like the variance and [3] for general certainty equivalents with complete observation. Applications to risk management can be found in [10]. In [7] the authors discuss the influence of the utility function on the optimal strategy.

Our paper is organized as follows. In the next section we present a short introduction to certainty equivalents and discuss their properties. In particular we show that they can serve as a reasonable criterion for risk-sensitive decision making. In Section 3 we introduce our risk-sensitive stopping problem with partial observation and with finite time horizon. We prove the existence of an optimal stopping time and give a recursive algorithm for the computation of the maximal expected utility. We are also able to show that the more risk averse a decision maker is, the later she will stop. In Section 4 we will consider the Bayesian, risk-sensitive house selling problem. Here we generalize the results in [17] to the Bayesian case. In particular we show the existence of so-called *reservation levels* which characterize the optimal stopping strategy. We also discuss the dependence of the reservation levels on the available information, the attitude towards risk of the decision maker and the time horizon. Finally in Section 5 we consider the general stopping problem with infinite time horizon.

## 2. CERTAINTY EQUIVALENTS

In this section we briefly recall the properties of our objective function and demonstrate that it is very well-suited for risk-sensitive optimization. For a deeper investigation we refer the reader to [16]. In what follows we work with the space of all bounded real-valued random variables  $\mathcal{L}^\infty$ . We interpret the outcome of a random variable  $X$  as the realization of a risky position. The event  $X < 0$  represents a loss whereas  $X > 0$  represents a gain. For a strictly increasing function  $U : \mathbb{R} \rightarrow \mathbb{R}$  we define the *certainty equivalent*  $\rho_U : \mathcal{L}^\infty \rightarrow \mathbb{R}$  by

$$\rho_U(X) := U^{-1}(\mathbb{E}[U(X)]).$$

The quantity is called certainty equivalent since  $U(\rho_U(X)) = \mathbb{E}[U(X)]$  which means that the utility of  $\rho_U(X)$  is the same as the expected utility of  $X$ . Obviously  $\rho_U$  is *law-invariant*, i.e.  $\rho_U(X) = \rho_U(Y)$  if  $X$  and  $Y$  have the same distribution. In what follows we summarize some properties of  $\rho_U$ . For this purpose we assume that  $U$  is *strictly increasing and concave*. Recall that for two real-valued random variable  $X$  and  $Y$  we say that

- $X \leq_{st} Y$  if and only if  $\mathbb{E}f(X) \leq \mathbb{E}f(Y)$  for all increasing  $f : \mathbb{R} \rightarrow \mathbb{R}$  for which the expectations exist,
- $X \leq_{cv} Y$  if and only if  $\mathbb{E}f(X) \leq \mathbb{E}f(Y)$  for all concave  $f : \mathbb{R} \rightarrow \mathbb{R}$  for which the expectations exist.

It is now easy to see that  $\rho_U$  is *monotone*, i.e.

- **(monotonicity)**  $X \leq_{st} Y$  implies that  $\rho_U(X) \leq \rho_U(Y)$ .

Obviously we also have

- **(constancy)**  $\rho_U(m) = m$  for any constant  $m \in \mathbb{R}$ .

Moreover, we have consistency w.r.t. the concave ordering. This property is also called *Schur-concavity* (see e.g. [9]), i.e.

- **(Schur-concavity)**  $X \leq_{cv} Y$  implies that  $\rho_U(X) \leq \rho_U(Y)$ .

Note that Schur-concavity is not the same as *concavity of a risk measure*  $\rho : \mathcal{L}^\infty \rightarrow \mathbb{R}$  which is defined by

$$\rho(\alpha X + (1 - \alpha)Y) \geq \alpha \rho(X) + (1 - \alpha)\rho(Y), \text{ for all } \alpha \in [0, 1] \text{ and r.v. } X, Y \in \mathcal{L}^\infty.$$

It has been shown in [1] that any concave risk measure on a non-atomic probability space is Schur-concave. However, only when  $U$  is linear or exponential, the certainty equivalent  $\rho_U$  is also concave (see [16]).

Hence maximizing the certainty equivalent means to prefer risky positions with higher reward and lower variance. In the classical risk neutral setting, the variance of a position does not enter the decision criterion. This may sometimes lead to a very risky behavior (as e.g. demonstrated in [3] for a casino game). Finally let us consider the following axioms which are often imposed on a risk measure  $\rho : \mathcal{L}^\infty \rightarrow \mathbb{R}$ :

- **(translation-invariance)**  $\rho(X + m) = \rho(X) + m$  for any constant  $m \in \mathbb{R}$ ,
- **(positive homogeneity)**  $\rho(\alpha X) = \alpha \rho(X)$  for any constant  $\alpha > 0$ .

It has been shown in [16] that  $\rho_U$  is translation-invariant essentially if and only if  $U$  is linear or exponential and  $\rho_U$  is positive homogeneous if and only if  $U$  is the logarithm or a power function (including the linear function).

More general there exists an axiomatic characterization of certainty equivalents which is known as *Nagumo-Kolmogorov-de Finetti theorem*. It states that a functional  $\rho : \mathcal{L}^\infty \rightarrow \mathbb{R}$  is a certainty equivalent if and only if it is monotone, law invariant, quasi-linear and has the constancy property. For an early proof see [12]. For a more recent treatment see [16].

Another motivation for certainty equivalents is the fact that they may be written (assuming enough regularity of  $U$ ) as in (1.1). Thus the certainty equivalent is approximately a weighted criterion of expectation and variance.

An important special case is obtained when

$$U(x) = \frac{1}{\gamma} e^{\gamma x}, \quad x \in \mathbb{R}, \quad \text{with } \gamma < 0$$

is an exponential function. In this case

$$\rho_U(X) = \frac{1}{\gamma} \ln \mathbb{E} e^{\gamma X}$$

is the *entropic risk measure*. In this case if  $X \sim \mathcal{N}(\mu, \sigma^2)$  then we obtain by direct calculation that  $\rho_U(X) = \mu + \frac{1}{2}\gamma\sigma^2$ . The entropic risk measure also has a dual representation given by

$$\rho_U(X) = \inf_{\mathbb{Q} \ll \mathbb{P}} \left\{ \mathbb{E}_{\mathbb{Q}}[X] - \frac{1}{\gamma} \mathbb{E}_{\mathbb{Q}} \left[ \ln \frac{d\mathbb{Q}}{d\mathbb{P}} \right] \right\},$$

where the infimum is attained at

$$\mathbb{Q}^*(dz) = \frac{e^{\gamma z} \mathbb{P}(dz)}{\int e^{\gamma y} \mathbb{P}(dy)}.$$

Thus we can interpret the optimization problem in this case as a game against nature where the second player (nature) chooses the probability measure.

### 3. RISK-SENSITIVE STOPPING PROBLEMS

We suppose that a *partially observable Markov Process* is given which we introduce as follows: We denote this process by  $(X_n, Y_n)_{n \in \mathbb{N}_0}$  and assume a Borel state space  $E_X \times E_Y$ . The  $x$ -component will be the *observable* part, the  $y$ -component *cannot be observed* by the controller. There is a stochastic transition kernel  $Q$  from  $E_X \times E_Y$  to  $E_X \times E_Y$  which determines the distribution of the new state pair given the current state. So  $Q(B|x, y)$  is the probability that the next state pair is in  $B \in \mathcal{B}(E_X \times E_Y)$ , given the current state is  $(x, y)$ . In what follows we assume that the transition kernel  $Q$  has a density  $q$  with respect to some  $\sigma$ -finite measures  $\lambda$  and  $\nu$ , i.e.

$$Q(B|x, y) = \int_B q(x', y'|x, y) \lambda(dx') \nu(dy'), \quad B \in \mathcal{B}(E_X \times E_Y).$$

For convenience we introduce the marginal transition kernel density by

$$q^X(x'|x, y) := \int_{E_Y} q(x', y'|x, y) \nu(dy').$$

We assume that the initial distribution  $Q_0$  of  $Y_0$  is known. For a fixed (observable) initial state  $x \in E_X$ , the initial distribution  $Q_0$  together with the transition probability  $Q$  define by

the Theorem of Ionescu Tulcea a probability measure  $\mathbb{P}_x$  on  $(E_X \times E_Y)^{N+1}$  endowed with the product  $\sigma$ -algebra. More precisely  $\mathbb{P}_{xy}$  is the probability measure given  $X_0 = x$  and  $Y_0 = y$  and we define  $\mathbb{P}_x(\cdot) := \int P_{xy}(\cdot) Q_0(dy)$ .

Next we have a measurable one-stage reward function  $c : E_X \rightarrow \mathbb{R}$  which depends only on the observable part of the process. Since  $c < 0$  is possible this could also be a cost. This reward (cost) is obtained as long as the process is not stopped. A reward  $g : E_X \rightarrow \mathbb{R}$  is obtained when the process is stopped. Stopping-times are taken w.r.t. the observable filtration  $(\mathcal{F}_n)$  where  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ . We first consider problems with a finite time horizon  $N$  and maximize the certainty equivalent of the stopping reward. Since  $U^{-1}$  is increasing we can skip it from the optimization problem. Thus, we define

$$J_N(x) := \sup_{0 \leq \tau \leq N} \mathbb{E}_x \left[ U \left( \sum_{k=0}^{\tau-1} c(X_k) + g(X_\tau) \right) \right]. \quad (3.1)$$

In order to have a well-defined problem we assume that for all  $x \in E_X$

$$\sup_{0 \leq \tau \leq N} \sup_{y \in E_Y} \mathbb{E}_{xy} \left[ \sum_{k=0}^{\tau-1} c^+(X_k) + g^+(X_\tau) \right] < \infty. \quad (3.2)$$

Note that since  $U$  is concave it can be bounded by a linear function and thus (3.2) implies that  $J_N(x) < \infty$  for all  $x \in E_X$ . Unfortunately (3.1) cannot be solved with classical MDP techniques directly. Thus, we introduce the following auxiliary problems. For a probability measure  $\mu \in \mathbb{P}(E_Y)$  and a constant  $s \in \mathbb{R}$  define for  $n = 1, \dots, N$ :

$$V_n(x, \mu, s) := \sup_{0 \leq \tau \leq n} \int_{E_Y} \mathbb{E}_{xy} \left[ U \left( \sum_{k=0}^{\tau-1} c(X_k) + s + g(X_\tau) \right) \right] \mu(dy) \quad (3.3)$$

In view of (3.2)  $V_n(x, \mu, s) < \infty$  for all  $(x, \mu, s)$ . Obviously we have by this embedding technique that  $J_N(x) = V_N(x, Q_0, 0)$ . We claim now that (3.3) can be solved with MDP techniques. In order to do so, we first have to cope with the fact that  $(Y_n)$  is not observable. Hence we define the operator  $\Phi : E_X \times E_X \times \mathbb{P}(E_Y) \rightarrow \mathbb{P}(E_Y)$  by

$$\Phi(x, x', \mu)(B) := \frac{\int_B q^X(x'|x, y) \mu(dy)}{\int_{E_Y} q^X(x'|x, y) \mu(dy)}, \quad B \in \mathcal{B}(E_Y).$$

Note that  $\Phi$  is exactly the usual updating (Bayesian) operator which appears in classical POMDP (see e.g. [4], Section 5.2). It updates the conditional probability of the unobservable state. In what follows denote by  $(\mu_n)$  the sequence of probability measures on  $E_Y$  generated by  $\Phi$  with  $\mu_0 := Q_0$ . I.e. for  $n \in \mathbb{N}$ , a history  $h_n := (x_0, \dots, x_n) \in E_X^{n+1}$  and  $B \in \mathcal{B}(E_Y)$  we define

$$\begin{aligned} \mu_0(B|h_0) &:= Q_0(B), \\ \mu_{n+1}(B|h_n, x') &:= \Phi(x_n, x', \mu_n(\cdot|h_n))(B) \end{aligned} \quad (3.4)$$

Then it is well-known (see e.g. [4], Theorem 5.2.1) that

$$\mu_n(B|X_0, \dots, X_n) = \mathbb{P}_x(Y_n \in B|X_0, \dots, X_n), \quad B \in \mathcal{B}(E_Y).$$

We consider now a stopping problem with a process on the state space  $E := E_X \times \mathbb{P}(E_Y) \times \mathbb{R}$ . The running reward is zero and the stopping reward is  $g(x, \mu, s) := U(g(x) + s)$ . The transition law of the process is given by  $\tilde{Q}(\cdot|x, \mu, s)$  which is for  $(x, \mu, s) \in E$  and a measurable function  $v : E \rightarrow \mathbb{R}$  defined by

$$\begin{aligned} & \int_{E_X} \int_{\mathbb{P}(E_Y)} \int_{\mathbb{R}} v(x', \mu', s') \tilde{Q}(d(x', \mu', s')|x, \mu, s) \\ &= \int_{E_X} v(x', \Phi(x, x', \mu), s + c(x)) Q^X(dx'|x, \mu) \end{aligned}$$

where

$$Q^X(B|x, \mu) := \int_B \int_{E_Y} q^X(x'|x, y) \mu(dy) \lambda(dx'). \quad (3.5)$$

Stopping times are now considered w.r.t. the filtration  $(\mathcal{G}_n)$  which is defined by  $\mathcal{G}_n = \sigma(X_0, \mu_0, S_0, \dots, X_n, \mu_n, S_n)$  with  $S_n := \sum_{k=0}^{n-1} c(X_k)$ . However note that obviously by construction of the sequences  $(\mu_n)$  and  $(S_n)$  we obtain  $\mathcal{F}_n = \mathcal{G}_n$ .

**Theorem 3.1.** *It holds:*

a) For  $n = 1, \dots, N$  and  $(x, \mu, s) \in E_X \times \mathbb{P}(E_Y) \times \mathbb{R}$ , the functions  $V_n$  are given by

$$\begin{aligned} V_0(x, \mu, s) &= U(g(x) + s) \\ V_n(x, \mu, s) &= \max \left\{ U(g(x) + s), \int_{E_X} V_{n-1}(x', \Phi(x, x', \mu), s + c(x)) Q^X(dx'|x, \mu) \right\} \end{aligned} \quad (3.6)$$

The value function of (3.1) is then given by  $J_N(x) = V_N(x, Q_0, 0)$ .

b) For every  $n = 1, \dots, N$  and  $(x, \mu, s) \in E_X \times \mathbb{P}(E_Y) \times \mathbb{R}$ , let  $f_n^*(x, \mu, s) = 1$  if the maximum in the recursion (3.6) is taken at  $U(g(x) + s)$  and define  $(g_0^*, \dots, g_{N-1}^*)$  by

$$g_n^*(h_n) := f_{N-n}^*(x_n, \mu_n(\cdot|h_n), \sum_{k=0}^{n-1} c(x_k)), \quad n = 0, \dots, N-1.$$

Then the optimal stopping time is given by

$$\tau^* := \inf\{n \in \mathbb{N}_0 : g_n^*(h_n) = 1\} \wedge N.$$

Note that  $\tau^*$  depends on the history  $h_n$  of the process.

*Proof.* The proof follows from Theorem 3.3 in [2].  $\square$

**Remark 3.2.** By definition we have  $U(g(x) + s) \leq V_n(x, \mu, s) \leq V_{n+1}(x, \mu, s)$  for all  $n \in \mathbb{N}$  and  $(x, \mu, s) \in E_X \times \mathbb{P}(E_Y) \times \mathbb{R}$  i.e. the value of the stopping problem increases with the time horizon. This observation is of course obvious, since all stopping times which are feasible for a smaller time horizon are also feasible for a larger time horizon. Note that we use increasing and decreasing in a non-strict sense throughout.

**Example 3.3.** Let  $U(x) = \frac{1}{\gamma} e^{\gamma x}$  with  $\gamma < 0$ . In this case Theorem 3.1 simplifies due to the positive homogeneity of  $\rho_U$  and it is easy to see that

$$V_n(x, \mu, s) = e^{\gamma s} h_n(x, \mu)$$

and the  $h_n$  satisfy the recursion

$$\begin{aligned} h_0(x, \mu) &= \frac{1}{\gamma} e^{\gamma g(x)} \\ h_n(x, \mu) &= \max \left\{ \frac{1}{\gamma} e^{\gamma g(x)}, e^{\gamma c(x)} \int_{E_X} h_{n-1}(x', \Phi(x, x', \mu)) Q^X(dx'|x, \mu) \right\}. \end{aligned}$$

As a result, in the exponential case the state space is reduced by one variable and the state of the auxiliary problem consists only of  $x$  and the conditional distribution of  $y$  given the history  $h_n$ . In particular the optimal stopping rule depends on the history only through the conditional probability of the unobservable random variable. The same situation arises in the classical risk-neutral stopping problem.

Finally we discuss the influence of the risk attitude of the decision maker on the optimal stopping time. We use the Arrow-Pratt function of absolute risk aversion (1.2) to measure the risk sensitivity. A utility function  $U$  is said to be *more risk averse* than a utility function  $W$  if  $l_U(x) \geq l_W(x)$  for all  $x \in \mathbb{R}$ . For our purpose it is crucial to note that a utility function  $U$  is more risk averse than a utility function  $W$  if and only if, there exists an increasing concave

function  $r : \mathbb{R} \rightarrow \mathbb{R}$  such that  $U = r \circ W$ . In what follows we denote all quantities which refer to utility function  $U$  by  $g_n^*(h_n, U)$ ,  $V_n(x, \mu, s, U)$  etc. and similar for  $W$ .

**Theorem 3.4.** *Suppose that the utility function  $U$  is more risk averse than the utility function  $W$ . For all  $n = 0, 1, \dots, N-1$  and histories  $h_n$  we obtain that  $g_n^*(h_n, W) = 1$  implies  $g_n^*(h_n, U) = 1$ , i.e. a more risk-averse decision maker will not stop later.*

*Proof.* Let  $r$  be such that  $U = r \circ W$ . We first prove by induction on  $n$  that  $V_n(x, \mu, s, U) \leq r \circ V_n(x, \mu, s, W)$  for all  $(x, \mu, s)$  and  $n$ . First for  $n = 0$  we have

$$V_0(x, \mu, s, U) = U(g(x) + s) = r \circ W(g(x) + s) = r \circ V_0(x, \mu, s, W).$$

Using the Jensen inequality, the induction hypothesis and the fact that  $r$  is increasing and concave we obtain

$$\begin{aligned} V_n(x, \mu, s, U) &= \\ &= \max \left\{ U(g(x) + s), \int_{E_X} V_{n-1}(x', \Phi(x, x', \mu), s + c(x), U) Q^X(dx'|x, \mu) \right\} \\ &\leq \max \left\{ r \circ W(g(x) + s), \int_{E_X} r \circ V_{n-1}(x', \Phi(x, x', \mu), s + c(x), W) Q^X(dx'|x, \mu) \right\} \\ &\leq \max \left\{ r \circ W(g(x) + s), r \circ \int_{E_X} V_{n-1}(x', \Phi(x, x', \mu), s + c(x), W) Q^X(dx'|x, \mu) \right\} \\ &= r \circ \max \left\{ W(g(x) + s), \int_{E_X} V_{n-1}(x', \Phi(x, x', \mu), s + c(x), W) Q^X(dx'|x, \mu) \right\} \\ &= r \circ V_n(x, \mu, s, W). \end{aligned}$$

This implies in particular that

$$W(g(x) + s) \geq \int_{E_X} V_{n-1}(x', \Phi(x, x', \mu), s + c(x), W) Q^X(dx'|x, \mu)$$

leads to

$$\begin{aligned} U(g(x) + s) &= \\ &= r \circ W(g(x) + s) \geq r \circ \int_{E_X} V_{n-1}(x', \Phi(x, x', \mu), s + c(x), W) Q^X(dx'|x, \mu) \\ &\geq \int_{E_X} r \circ V_{n-1}(x', \Phi(x, x', \mu), s + c(x), W) Q^X(dx'|x, \mu) \\ &\geq \int_{E_X} V_{n-1}(x', \Phi(x, x', \mu), s + c(x), U) Q^X(dx'|x, \mu). \end{aligned}$$

By definition this means that  $f_n^*(x, \mu, s, W) = 1$  implies  $f_n^*(x, \mu, s, U) = 1$ . Transferring this observation to  $g_n^*$  implies the statement.  $\square$

#### 4. RISK-SENSITIVE BAYESIAN HOUSE SELLING PROBLEM

Suppose someone wants to sell her house. We assume that offers  $X_0, \dots, X_N$  arrive independently and are identically distributed with distribution  $Q_\theta$ . Here  $\theta \in \Theta$  is an unknown parameter. We assume that  $Q_\theta$  has a density  $q(x|\theta)$  and some prior distribution  $Q_0$  for  $\theta$  is given. As long as offers are rejected a maintenance cost of  $c > 0$  has to be paid. When an offer is accepted, the price is obtained and the game ends. The aim is to find the maximal risk-sensitive stopping reward

$$J_N(x) := \sup_{0 \leq \tau \leq N} \int_{\Theta} \mathbb{E}_{x\theta} \left[ U(X_\tau - c\tau) \right] Q_0(d\theta). \quad (4.1)$$

This is a special case of our general model with  $Y_n \equiv \theta$ ,  $c(x) = -c$  and  $g(x) = x$ . The integrability condition (3.2) reduces to

$$\sup_{\theta} \mathbb{E}_{\theta}[X_1^+] < \infty.$$

Note that this problem without partial observation has been investigated in [17]. According to Theorem 3.1 we obtain  $J_N$  by computing the functions  $V_n$ . These are given by

$$\begin{aligned} V_0(x, \mu, s) &= U(x + s) \\ V_n(x, \mu, s) &= \max \left\{ U(x + s), c_n(\mu, s) \right\} \end{aligned}$$

with  $c_n(\mu, s) := \int_{\mathbb{R}} V_{n-1}(x', \Phi(x', \mu), s - c) Q(dx' | \mu)$  and we have that  $J_N(x) = V_N(x, Q_0, 0)$ . Note that  $Q(dx' | \mu)$  corresponds to  $Q^X(dx' | x, \mu)$  in (3.5) and is thus given by

$$Q(\cdot | \mu) = \int_{\Theta} Q_{\theta}(\cdot) \mu(d\theta)$$

and the updating operator simplifies to

$$\Phi(x, \mu)(B) := \frac{\int_B q(x | \theta) \mu(d\theta)}{\int_{\Theta} q(x | \theta) \mu(d\theta)}, \quad B \in \mathcal{B}(\Theta).$$

Moreover, when we define  $f_n^*(x, \mu, s) = 1$  if  $U(x + s) \geq c_n(\mu, s)$  and  $(g_0^*, \dots, g_{N-1}^*)$  by

$$g_n^*(h_n) := f_{N-n}^*(x_n, \mu_n(\cdot | h_n), -nc), \quad n = 0, \dots, N-1,$$

then the optimal stopping time for problem (4.1) is given by

$$\tau^* := \inf \{n \in \mathbb{N}_0 : g_n^*(h_n) = 1\} \wedge N.$$

Let us now further investigate the optimal stopping time. We have

$$g_n^*(h_n) = 1 \quad \Leftrightarrow \quad U(x_n - nc) \geq c_{N-n}(\mu_n(\cdot | h_n), -nc).$$

When we define

$$\begin{aligned} U_n(x) &:= U(x - nc) \\ d_k(\mu, U_n) &:= c_k(\mu, -nc) = \int_{\mathbb{R}} V_{k-1}(x', \Phi(x', \mu), -(n+1)c) Q(dx' | \mu), \end{aligned}$$

then we obtain

$$g_n^*(h_n) = 1 \quad \Leftrightarrow \quad x_n \geq U_n^{-1}(d_{N-n}(\mu_n(\cdot | h_n), U_n)) =: x_{n,N}^*(\mu_n(\cdot | h_n)).$$

Note that  $U_n^{-1}(x) = nc + U^{-1}(x)$ . We call  $x_{n,N}^*(\cdot)$  *reservation level*. The reservation levels depend on  $\mu_n$  and  $U$ . The optimal stopping time is hence the first time, the offer exceeds the corresponding, history dependend reservation level.

**Theorem 4.1.** a) *The optimal stopping time for the Bayesian house selling problem is given by*

$$\tau^* = \inf \{n \in \mathbb{N}_0 : X_n \geq x_{n,N}^*(\mu_n(\cdot | h_n))\} \wedge N.$$

b) *The reservation levels can recursively be computed by*

$$\begin{aligned} x_{N-1,N}^*(\mu) &= U_{N-1}^{-1} \circ \int_{\mathbb{R}} U_N(x) Q(dx | \mu) \\ x_{n,N}^*(\mu) &= U_n^{-1} \circ \int_{\mathbb{R}} U_{n+1} \left( \max \{x, x_{n+1,N}^*(\Phi(x, \mu))\} \right) Q(dx | \mu). \quad \square \end{aligned}$$

*Proof.* Part a) is clear from the definition and the previous results. Part b) can be shown by inserting the correct definitions. For  $n = N - 1$  we obtain from the definition of  $x_{N-1,N}^*(\mu)$  that

$$x_{N-1,N}^*(\mu) = U_{N-1}^{-1}(d_1(\mu, U_{N-1}))$$

with

$$d_1(\mu, U_{N-1}) = \int V_0(x, \Phi(x, \mu), -Nc)Q(dx|\mu) = \int U_N(x)Q(dx|\mu).$$

For  $x_{n,N}^*$  we obtain by definition:

$$x_{n,N}^*(\mu) = U_n^{-1}(d_{N-n}(\mu, U_n)).$$

Further  $d_{N-n}(\mu, U_n)$  can be written as

$$\begin{aligned} d_{N-n}(\mu, U_n) &= \int_{\mathbb{R}} V_{N-n-1}\left(x, \Phi(x, \mu), -(n+1)c\right)Q(dx|\mu) \\ &= \int_{\mathbb{R}} \max\left\{U(x - (n+1)c), \right. \\ &\quad \left. \int_{\mathbb{R}} V_{N-n-2}\left(x', \Phi(x', \Phi(x, \mu)), -(n+2)c\right)Q(dx'|\Phi(x, \mu))\right\}Q(dx|\mu) \\ &= \int_{\mathbb{R}} \max\left\{U_{n+1}(x), d_{N-n-1}(\Phi(x, \mu), U_{n+1})\right\}Q(dx|\mu) \\ &= \int_{\mathbb{R}} U_{n+1}\left(\max\left\{x, U_{n+1}^{-1} \circ d_{N-n-1}(\Phi(x, \mu), U_{n+1})\right\}\right)Q(dx|\mu) \end{aligned}$$

and the statement follows from the definition of  $x_{n+1,N}^*$ .  $\square$

**Remark 4.2.** It obviously holds that

$$\begin{aligned} d_N(\mu, U) &= \int_{\mathbb{R}} V_{N-1}\left(x, \Phi(x, \mu), -c\right)Q(dx|\mu) \\ &= \sup_{1 \leq \tau \leq N} \int_{\Theta} \int_{\mathbb{R}} \mathbb{E}_{x\theta}[U(X_\tau - c\tau)]Q_\theta(dx)\mu(d\theta), \end{aligned}$$

i.e.  $d_N(Q_0, U)$  is the value of the stopping problem when we start without known initial offer.

**Example 4.3.** In case  $U(x) = \frac{1}{\gamma}e^{\gamma x}$  with  $\gamma < 0$  the recursion for the reservation levels simplifies. In order to see this, note that  $U_n(x) = U(x - nc) = \frac{1}{\gamma}e^{\gamma x} \cdot e^{-\gamma nc}$  and  $U_n^{-1}(x) = U^{-1}(x) + nc = \frac{1}{\gamma} \ln(\gamma x) + nc$ . With these observation we obtain

$$\begin{aligned} x_{N-1,N}^*(\mu) &= -c + \frac{1}{\gamma} \ln \left( \int_{\mathbb{R}} e^{\gamma x} Q(dx|\mu) \right) \\ x_{n,N}^*(\mu) &= -c + \frac{1}{\gamma} \ln \left( \int_{\mathbb{R}} e^{\gamma \max\{x, x_{n+1,N}^*(\Phi(\mu, x))\}} Q(dx|\mu) \right). \end{aligned}$$

Note that in contrast to the  $x_{n,N}^*$  with general utility function, the reservation levels in the exponential utility case depend only on the time difference to the planning horizon. More precisely we could also define  $x_n^*(\mu) := x_{N-n,N}^*(\mu)$  and obtain

$$\begin{aligned} x_1^*(\mu) &= -c + \frac{1}{\gamma} \ln \left( \int_{\mathbb{R}} e^{\gamma x} Q(dx|\mu) \right) \\ x_n^*(\mu) &= -c + \frac{1}{\gamma} \ln \left( \int_{\mathbb{R}} e^{\gamma \max\{x, x_{n-1}^*(\Phi(\mu, x))\}} Q(dx|\mu) \right). \end{aligned}$$

In particular we can start computing the reservation levels without fixing a planning horizon in advance.

We did this in the following *numerical example*, where  $Q_\theta = B(1, \theta)$  is a Bernoulli distribution with unknown 'success' probability  $\theta$ , the prior distribution of  $\theta$  is uniform on  $[0, 1]$  and  $c = 0.1$ . For a planning horizon of  $N = 10$  we computed the optimal stopping rule. Of course an offer



$-\infty < \gamma < -2.2$	$-2.2 < \gamma < -1.51$	$-1.51 < \gamma < -1.1$	$-1.1 < \gamma < -0.8$
0	1	2	3
$-0.8 < \gamma < -0.56$	$-0.56 < \gamma < -0.34$	$-0.34 < \gamma < -0.18$	$-0.18 < \gamma < -0.03$
4	5	6	7

FIGURE 1. Optimal number of zeros which are rejected.

of 1 will always be accepted. An offer of 0 may be accepted when we fear the accumulation of cost. It turns out that this decision heavily depends on the risk aversion parameter  $\gamma < 0$ . The smaller  $\gamma$ , the more risk averse the decision maker is. **Figure 1** has to be read as follows: When  $\gamma < -2.2$ , the decision maker will stop immediately. For  $-2.2 < \gamma < -1.51$  she will reject at least the first offer when it is zero. For  $-1.51 < \gamma < -1.1$  she will reject the first two offers when they are zero. And so on. The switch from rejecting 8 to 9 zeros is for  $\gamma$  smaller than  $-10^{-8}$ .

**4.1. Influence of the Filter on the Reservation Levels.** In order to discuss the influence of the filter on the reservation levels we make some further simplifying assumptions. Indeed it is often the case that the filter  $\mu_n(\cdot|h_n)$  does only depend on a part of the history  $h_n$  or on a certain function of it. In what follows we assume that there is a (Borel) information set  $I$  endowed with a  $\sigma$ -algebra such that there exist a measurable function  $t_n : H_n \rightarrow I$  and a transition kernel  $\hat{\mu}$  from  $I$  to  $\Theta$  such that

$$\mu_n(B|h_n) = \hat{\mu}(B|t_n(h_n)), \quad \text{for } B \in \mathcal{B}(E_Y).$$

The function  $t_n$  is sometimes called *sufficient statistics*. Further we assume that there is a measurable mapping  $\hat{\Phi} : I \times E_X$  such that

$$t_{n+1}(h_{n+1}) = \hat{\Phi}(x_{n+1}, t_n(h_n))$$

and that  $\hat{\mu}$  has a density  $\hat{p}(\cdot|i)$ . Thus, we have to replace  $\mu_n$  by the current information state and obtain in particular for the reservation levels:

$$\begin{aligned} x_{N-1,N}^*(i) &:= U_{N-1}^{-1} \circ \int_{\mathbb{R}} U_N(x) Q(dx|i) \\ x_{n,N}^*(i) &:= U_n^{-1} \circ \int_{\mathbb{R}} U_{n+1} \left( \max \{x, x_{n+1,N}^*(\hat{\Phi}(x, i))\} \right) Q(dx|i) \end{aligned}$$

where

$$Q(B|i) := \int_B \int_{\Theta} q(x|\theta) \hat{\mu}(d\theta|i) \lambda(dx).$$

The next step is to introduce an order relation on the set  $I$  where we assume now that  $\Theta \subset \mathbb{R}$ . We define here for  $i, i' \in I$

$$i \leq i' \quad :\Leftrightarrow \quad \hat{\mu}(\cdot|i) \leq_{lr} \hat{\mu}(\cdot|i')$$

where  $\leq_{lr}$  is the likelihood ratio ordering which is defined by

$$\hat{\mu}(\cdot|i) \leq_{lr} \hat{\mu}(\cdot|i') \quad \Leftrightarrow \quad \frac{\hat{p}(\theta|i')}{\hat{p}(\theta|i)} \quad \text{is increasing in } \theta.$$

Note that  $\hat{p}$  is the density of  $\hat{\mu}$ . The likelihood ratio ordering implies the stochastic ordering. Now we are able to formulate the main result of this subsection

**Theorem 4.4.** *Suppose that  $q(x|\theta)$  is MTP<sub>2</sub>, i.e.  $q(\cdot|\theta) \leq_{lr} q(\cdot|\theta')$  for all  $\theta \leq \theta'$ , then the reservation levels  $x_{n,N}^*(i)$  are increasing in  $i$ .*

*Proof.* We prove the statement by induction on  $n$ . First consider

$$x_{N-1,N}^*(i) = U_{N-1}^{-1} \circ \int_{\mathbb{R}} U_N(x) \int_{\Theta} q(x|\theta) \hat{\mu}(d\theta|i) \lambda(dx).$$

Since  $x \mapsto U_N(x)$  is increasing and the  $\leq_{lr}$  implies the  $\leq_{st}$  order we obtain that

$$\theta \mapsto \int_{\mathbb{R}} U_N(x) q(x|\theta) \lambda(dx) =: f(\theta)$$

is increasing in  $\theta$ . Thus by the definition of  $i \leq i'$  we obtain that  $\int_{\Theta} f(\theta) \hat{\mu}(d\theta|i)$  is increasing in  $i$ . Now suppose the statement is true for  $x_{n+1,N}^*(i)$ . Next note that by our assumption on  $q(x|\theta)$  we have that

$$(i, x) \mapsto \hat{\Phi}(x, i)$$

is increasing. This follows from Lemma 5.4.9 in [4]. Hence

$$(i, x) \mapsto U_{n+1} \left( \max \{x, x_{n+1,N}^*(\hat{\Phi}(x, i))\} \right) =: f(i, x)$$

is increasing. Next by our assumption on  $q(x|\theta)$  we have that

$$(i, \theta) \mapsto \int_{\mathbb{R}} f(i, x) q(x|\theta) \lambda(dx) =: \hat{f}(i, \theta)$$

is increasing. And finally we obtain that

$$x_{n,N}^*(i) = U_n^{-1} \circ \int_{\Theta} \hat{f}(i, \theta) \hat{\mu}(d\theta|i)$$

is increasing which completes the induction.  $\square$

For further details and examples we refer the reader to Section 5.4 in [4].

**Example 4.5.** In this example we consider the special case of exponentially distributed random variables (offers)

$$q(x|\theta) = \frac{1}{\theta} e^{-\frac{1}{\theta}x}, \quad x \geq 0, \quad \theta \in \Theta := (0, \infty).$$

According to our definition of  $\mu_n$  we get by recursion for  $h_n = (x_1, \dots, x_n)$  and  $B \in \mathcal{B}(\Theta)$

$$\mu_n(B|h_n) = \frac{\int_B \frac{1}{\theta^n} \exp(-\frac{1}{\theta} \sum_{k=1}^n x_k) Q_0(d\theta)}{\int_{\Theta} \frac{1}{\theta^n} \exp(-\frac{1}{\theta} \sum_{k=1}^n x_k) Q_0(d\theta)}.$$

Thus  $t_n(x_1, \dots, x_n) = \left( \sum_{k=1}^n x_k, n \right)$  is a sufficient statistic. Thus, we have  $I := \mathbb{R}_+ \times \mathbb{N}_0$  and denote  $i = (s, n) \in I$ . Moreover,  $\hat{\Phi}(x, (s, n)) = (s + x, n + 1)$  and the conditional distribution of the unknown parameter has the form

$$\hat{\mu}(d\theta|s, n) \propto \left( \frac{1}{\theta} \right)^n e^{-\frac{s}{\theta}} Q_0(d\theta)$$

if the information  $(s, n)$  is given. With this representation it is not difficult to verify that

$$i = (s, n) \leq i' = (s', n') \Leftrightarrow s \leq s' \text{ and } n \geq n'.$$

Further, the family of densities  $q(x|\theta)$  is  $MTP_2$  in  $x \geq 0$  and  $\theta \geq 0$ , thus Theorem 4.4 applies.

If we assume now a special prior distribution of the unknown parameter then we obtain an explicit distribution for  $Q(\cdot|i)$ . We assume that the prior distribution  $Q_0$  is a so-called *Inverse Gamma distribution*, i.e. the density is given by

$$Q_0(d\theta) = \frac{b^a}{\Gamma(a)} \left( \frac{1}{\theta} \right)^{a+1} e^{-\frac{b}{\theta}} d\theta, \quad \theta > 0$$

where  $a > 1$  and  $b > 0$  are fixed. Then the distribution  $Q$  is given by

$$Q(dx|s, n) = \int q(dx|\theta) \hat{\mu}(d\theta|s, n) = (n + a) \frac{(s + b)^{n+a}}{(x + s + b)^{n+a+1}} dx.$$

Hence  $Q$  is a special *Second Order Beta distribution*. In a risk-neutral situation a similar setting has been considered in [22].

**4.2. Influence of the Utility Function on the Reservation Levels.** Here we proceed as in [17] (Theorem 3.3) in order to study the impact of the risk aversion on the reservation levels. As in section 3 we use the Arrow-Pratt function of absolute risk aversion (1.2) to measure the risk attitude in the sense that a utility function  $U$  is more risk averse than a utility function  $W$  if  $l_U(x) \geq l_W(x)$  for all  $x \in \mathbb{R}$ . In what follows we denote by  $x_{n,N}^*(\mu, U)$  the reservation levels which belong to the utility function  $U$ . Then we obtain

**Theorem 4.6.** *If  $U$  is a more risk averse utility function than  $W$ , then the reservation levels satisfy  $x_{n,N}^*(\mu, U) \leq x_{n,N}^*(\mu, W)$  for all  $n = 0, 1, \dots, N-1$  and all  $\mu \in \mathbb{P}(E_Y)$ .*

*Proof.* The proof follows from Theorem 3.4.  $\square$

**Remark 4.7.** Theorem 4.6 includes as a special case the comparison to the risk neutral stopping problem: Suppose  $U$  is an arbitrary increasing concave utility function, then obviously  $U = U \circ \text{id}$ . Thus, we choose  $r = U$  in this context and see that the reservation levels of a risk neutral decision maker will always be above the reservation levels of a risk averse decision maker.

**Example 4.8.** In case  $U(x) = \frac{1}{\gamma}e^{\gamma x}$  with  $\gamma < 0$  we obtain from Theorem 4.6 that the reservation levels  $x_n^*(\mu)$  are increasing in  $\gamma$ .

**4.3. Influence of the Time Horizon on the Reservation Levels.** In a risk neutral setting it is often the case that the reservation levels are decreasing as time goes by, i.e. the decision maker becomes less selective when she approaches the time horizon. However in [17] it has been shown that this is no longer true in a risk averse setting. Indeed he gave some examples where the reservation levels are increasing.

Without additional assumptions it is difficult to determine how the reservation levels behave in time. However in general the reservation levels satisfy the following relation.

**Theorem 4.9.** *For all  $n = 0, 1, \dots, N-2$  and  $\mu \in \mathbb{P}(E_Y)$  it holds that*

$$x_{n,N}^*(\mu) \geq nc + \rho_U \left( x_{n+1,N}^*(\Phi(X_1, \mu)) - (n+1)c \right).$$

*In case  $U(x) = \frac{1}{\gamma}e^{\gamma x}$  with  $\gamma < 0$  we obtain*

$$x_{n,N}^*(\mu) + c \geq \rho_U \left( x_{n+1,N}^*(\Phi(X_1, \mu)) \right).$$

*Proof.* From Theorem 4.1 b) we obtain

$$\begin{aligned} x_{n,N}^*(\mu) &= U_n^{-1} \circ \int_{\mathbb{R}} U_{n+1} \left( \max \{x, x_{n+1,N}^*(\Phi(x, \mu))\} \right) Q(dx|\mu) \\ &\geq U_n^{-1} \circ \int_{\mathbb{R}} U_{n+1} \left( x_{n+1,N}^*(\Phi(x, \mu)) \right) Q(dx|\mu) \\ &= nc + U^{-1} \circ \int_{\mathbb{R}} U \left( x_{n+1,N}^*(\Phi(x, \mu)) - (n+1)c \right) Q(dx|\mu). \end{aligned}$$

Using the definition of the certainty equivalent yields the first statement. For the second statement note that  $\rho$  is translation invariant in the case of exponential utility.  $\square$

In order to obtain decreasing reservation levels further assumptions are necessary. For example the property that the utility function has a *decreasing absolute risk aversion* (DARA). This means that  $x \mapsto l_U(x)$  is decreasing, i.e. the decision maker becomes more risk averse with decreasing wealth.

**Theorem 4.10.** a) *The reservation levels  $x_{n,N}^*(\mu)$  are increasing in  $N$  for all  $\mu \in \mathbb{P}(E_Y)$ .*  
b) *If the utility function  $U$  is DARA, then  $x_{n,N}^*(\mu)$  are decreasing in  $n$  for all  $\mu \in \mathbb{P}(E_Y)$ .*

*Proof.* a) It is obvious that  $x_{N-1,N}^*$  is increasing in  $N$ . Now suppose that  $x_{n+1,N}^*$  is increasing in  $N$ . Then due to the recursion of the reservation levels in Theorem 4.1 and the fact that  $U$  is increasing we obtain that  $x_{n,N}^*$  is increasing in  $N$ .  
 b) Now suppose the utility function  $U$  is DARA. Thus, in particular  $U_1(x) = U(x - c)$  is more risk averse than  $U$ . Then we obtain with Theorem 4.6 that

$$x_{n,N}^*(\mu, U) \geq x_{n,N}^*(\mu, U_1) = x_{n+1,N}^*(\mu)$$

which implies the result.  $\square$

**Example 4.11.** In case  $U(x) = \frac{1}{\gamma}e^{\gamma x}$  with  $\gamma < 0$  we obtain from Theorem 4.10 that the reservation levels  $x_n^*(\mu)$  are increasing in  $n$ . This effect can also be seen in the numerical example 4.3.

## 5. RISK-SENSITIVE STOPPING PROBLEMS WITH INFINITE TIME HORIZON

Let us now consider the risk-sensitive stopping problem from Section 3 with infinite time horizon. Here we assume that the stopping reward  $g$  is bounded, i.e.  $\underline{g} \leq g \leq \bar{g}$  and cost are strictly negative, i.e.  $\sup_{x \in E_X} c(x) =: \bar{c} < 0$ . Thus we consider

$$J_\infty(x) := \sup_{\tau < \infty} \mathbb{E}_x \left[ U \left( \sum_{k=0}^{\tau-1} c(X_k) + g(X_\tau) \right) \right] \quad (5.1)$$

where the supremum is taken over all  $(\mathcal{F}_n)$ -stopping times  $\tau$  with  $\mathbb{P}_x(\tau < \infty) = 1$  for all  $x \in E_X$ . This problem can be seen as the limiting problem of stopping problems with bounded horizon.

**Theorem 5.1.** *The sequence  $(V_n)$  of value functions defined in (3.3) has a limit  $V$  and this limit satisfies*

$$V(x, \mu, s) = \max \left\{ U(g(x) + s), \int_{E_X} V(x', \Phi(x, x', \mu), s + c(x)) Q(dx' | \mu) \right\}. \quad (5.2)$$

*Proof.* We have  $V_n \leq V_{n+1}$  and the sequence  $V_n$  is bounded from above by our assumptions, thus the limit  $V = \lim_{n \rightarrow \infty} V_n$  exists. Moreover, we can take the limit on both sides in the recursion (3.6) for  $V_n$  to obtain the fixed point property of  $V$ .  $\square$

Next we show the relation between  $J_\infty$  and  $V$ .

**Theorem 5.2.** a) *It holds that  $V(x, Q_0, 0) = J_\infty(x)$ .*

b) *Let  $f^*(x, \mu, s) = 1$  if the maximum in (5.2) is attained at  $U(g(x) + s)$  and define  $(g_0^*, g_1^* \dots)$  by*

$$g_n^*(h_n) := f^* \left( x_n, \mu_n(\cdot | h_n), \sum_{k=0}^{n-1} c(x_k) \right), \quad n \in \mathbb{N}_0.$$

*Then the optimal stopping time  $\tau^*$  is given by*

$$\tau^* := \inf \{ n \in \mathbb{N}_0 : g_n^*(h_n) = 1 \}.$$

*Proof.* a) First note that we have  $V_n(x, Q_0, 0) \leq J_\infty(x)$  for all  $n \in \mathbb{N}$  which implies that  $V(x, Q_0, 0) \leq J_\infty(x)$ . Next for every admissible stopping time  $\tau$  with  $\mathbb{P}_x(\tau < \infty) = 1$

and  $\mathbb{E}_x \left[ \left| U \left( \sum_{k=0}^{\tau-1} c(X_k) + g(X_\tau) \right) \right| \right] < \infty$  it holds for any  $n \in \mathbb{N}$  that

$$\begin{aligned} V(x, Q_0, 0) &\geq V_n(x, Q_0, 0) \geq \\ &\geq \mathbb{E}_x \left[ U \left( \sum_{k=0}^{(\tau \wedge n)-1} c(X_k) + g(X_{\tau \wedge n}) \right) \right] \\ &\geq \mathbb{E}_x \left[ U \left( \sum_{k=0}^{\tau-1} c(X_k) + g(X_{\tau \wedge n}) \right) \right] \\ &= \mathbb{E}_x \left[ U \left( \sum_{k=0}^{\tau-1} c(X_k) + g(X_\tau) \right) 1_{[\tau \leq n]} \right] + \mathbb{E}_x \left[ U \left( \sum_{k=0}^{\tau-1} c(X_k) + g(X_n) \right) 1_{[\tau > n]} \right] \end{aligned}$$

Then letting  $n \rightarrow \infty$  yields with dominated convergence and the fact that  $\mathbb{P}_x(\tau < \infty) = 1$

$$V(x, Q_0, 0) \geq \mathbb{E}_x \left[ U \left( \sum_{k=0}^{\tau-1} c(X_k) + g(X_\tau) \right) \right].$$

Taking the supremum over all admissible stopping times and combining the result with the first inequality implies the result.

b) Iterating the fixed point equation  $n$ -times and using the definition of  $\tau^*$  we obtain

$$\begin{aligned} V(x, Q_0, 0) &= \mathbb{E}_x \left[ U \left( \sum_{k=0}^{\tau^*-1} c(X_k) + g(X_{\tau^*}) \right) 1_{[\tau^* \leq n]} \right] + \\ &\quad + \mathbb{E}_x \left[ V \left( X_n, \mu_n(\cdot | H_n), \sum_{k=0}^{n-1} c(X_k) \right) 1_{[\tau^* > n]} \right]. \end{aligned} \quad (5.3)$$

First we show that  $\mathbb{P}_x(\tau^* < \infty) = 1$  for all  $x \in E_X$ . We can extend part a) easily to arbitrary states  $(x, \mu, s)$ , i.e.

$$V(x, \mu, s) = \sup_{\tau < \infty} \int_{E_Y} \mathbb{E}_{xy} \left[ U \left( \sum_{k=0}^{\tau-1} c(X_k) + g(X_\tau) + s \right) \right] \mu(dy).$$

Then we obtain

$$V \left( x_n, \mu_n(\cdot | h_n), \sum_{k=0}^{n-1} c(x_k) \right) \leq U(n\bar{c} + \bar{g})$$

and from (5.3)

$$\begin{aligned} U(g(x)) &\leq V(x, Q_0, 0) \leq \\ &\leq U(\bar{g}) \mathbb{P}_x(\tau^* \leq n) + U(n\bar{c} + \bar{g}) \mathbb{P}_x(\tau^* > n), \end{aligned}$$

i.e.

$$\begin{aligned} U(g(x)) &\leq U(\bar{g}) + \left( U(n\bar{c} + \bar{g}) - U(\bar{g}) \right) \mathbb{P}_x(\tau^* > n) \\ &= U(\bar{g}) + a_n \mathbb{P}_x(\tau^* > n) \end{aligned}$$

where  $a_n := U(n\bar{c} + \bar{g}) - U(\bar{g})$ . It holds that  $a_n < 0$  for all  $n \in \mathbb{N}$  large enough and  $\lim_{n \rightarrow \infty} a_n = -\infty$ . Hence in total

$$\mathbb{P}_x(\tau^* > n) \leq \frac{U(g(x)) - U(\bar{g})}{a_n}$$

for all  $n \in \mathbb{N}$  large enough which implies  $\mathbb{P}_x(\tau^* < \infty) = 1$ . Since  $V$  is bounded from above, letting  $n \rightarrow \infty$  in (5.3) implies that

$$J_\infty(x) = V(x, Q_0, 0) \leq \mathbb{E}_x \left[ U \left( \sum_{k=0}^{\tau^*-1} c(X_k) + g(X_{\tau^*}) \right) \right]$$

and  $\tau^*$  is optimal. □

**5.1. Risk-Sensitive Bayesian House Selling Problem.** Let us consider the risk-sensitive Bayesian house selling problem with infinite time horizon. When we assume that the offers are bounded, i.e.  $X_i \in [m, M]$  our general assumptions of this section are satisfied. As in the finite horizon case we can see from the fixed point equation that the optimal stopping time is characterized by *reservation levels*. Indeed Theorem 5.1 and Theorem 5.2 apply directly and the fixed point equation reads

$$V(x, \mu, s) = \max \left\{ U(x + s), \int_{E_X} V(x', \Phi(x', \mu), s - c) Q(dx' | \mu) \right\}.$$

Then we obtain

$$\begin{aligned} g_n^*(h_n) = 1 & \Leftrightarrow x_n \geq U_n^{-1} \left( \int V(x', \Phi(x', \mu), -c(n+1)) Q(dx' | \mu) \right) \\ & =: x_{n,\infty}^*(\mu). \end{aligned}$$

Note that the reservation levels in general still depend on the time stage, in contrast to the problem in the risk neutral setting because we have to memorize the cost which has accumulated so far. As for the case of finite time horizon we obtain a similar recursion for the reservation levels.

**Theorem 5.3.** a) *The optimal stopping time for the Bayesian house selling problem with infinite time horizon is given by*

$$\tau^* = \inf \{ n \in \mathbb{N}_0 : X_n \geq x_{n,\infty}^*(\mu_n(\cdot | h_n)) \}.$$

b) *It holds that  $x_{n,\infty}^*(\mu) = \lim_{N \rightarrow \infty} x_{n,N}^*(\mu)$  and the reservation levels satisfy the following recursion*

$$x_{n,\infty}^*(\mu) = U_n^{-1} \circ \int_{\mathbb{R}} U_{n+1} \left( \max \{ x, x_{n+1,\infty}^*(\Phi(x, \mu)) \} \right) Q(dx | \mu).$$

**Example 5.4.** In case  $U(x) = \frac{1}{\gamma} e^{\gamma x}$  with  $\gamma < 0$  the recursion for the reservation levels simplifies to a fixed point equation

$$x_{\infty}^*(\mu) = -c + \frac{1}{\gamma} \ln \left( \int_{\mathbb{R}} e^{\gamma \max \{ x, x_{\infty}^*(\Phi(x, \mu)) \}} Q(dx | \mu) \right).$$

Moreover, it holds  $x_{\infty}^*(\mu) = \lim_{n \rightarrow \infty} x_n^*(\mu)$ . From Example 4.8 it follows that the reservation levels  $x_{\infty}^*(\mu)$  are increasing in  $\gamma$ .

Also it is possible to discuss the influence of the attitude towards risk as before. Here we use the notation  $x_{n,\infty}^*(\mu, U)$  when the reservation level belongs to utility function  $U$ .

**Theorem 5.5.** a) *Suppose that utility function  $U$  is more risk averse than utility function  $W$ . Then for all  $n \in \mathbb{N}$  we obtain that  $x_{n,\infty}^*(\mu, U) \leq x_{n,\infty}^*(\mu, W)$ .*

b) *If the utility function  $U$  is DARA, then  $x_{n,\infty}^*(\mu)$  is decreasing in  $n$ .*

*Proof.* For part a) we proceed as in the proof of Theorem 4.6 and show first that  $V_n(x, \mu, s, z, U) \leq r \circ V_n(x, \mu, s, z, W)$  for  $n \in \mathbb{N}$  where  $r$  is such that  $U = r \circ W$ . Taking the limit yields  $V(x, \mu, s, z, U) \leq r \circ V(x, \mu, s, z, W)$  and the statement follows as in the proof of Theorem 4.6. Part b) follows from part a). □

## 6. CONCLUSION

We have seen that the theory for partially observable risk-sensitive stopping problems is only slightly more complicated than the theory for the risk neutral case. We were also able to show in general that the more risk averse a decision maker is, the later she will stop. Though the numerical algorithms are more demanding than in the risk neutral case, the setting with exponential utility is still feasible.

## REFERENCES

- [1] Bäuerle, N. and Müller, A.: Stochastic orders and risk measures: consistency and bounds. *Insurance: Mathematics and Economics*, **38** (2006) 132-148.
- [2] Bäuerle, N. and Rieder, U. *Partially Observable Risk-Sensitive Markov Decision Processes*. Preprint (2015).
- [3] Bäuerle, N. and Rieder, U.: More risk-sensitive Markov decision processes. *Mathematics of Operations Research*, **39** (1) (2014) 105-120.
- [4] Bäuerle, N. and Rieder, U. *Markov Decision Processes with Applications to Finance*. Springer-Verlag, Berlin Heidelberg, 2011.
- [5] Bielecki, T. and Pliska, S.: Economic properties of the risk sensitive criterion for portfolio management. *Review of Accounting and Finance*, **2** (2003) 3-17.
- [6] Cavazos-Cadena, R. and Hernández-Hernández, D.: Successive approximations in partially observable controlled Markov chains with risk-sensitive average criterion. *Stochastics*, **77** (2005) 537-568.
- [7] Cavazos-Cadena, R. and Hernández-Hernández, D.: A Characterization of the Optimal Certainty Equivalent of the Average Cost via the Arrow-Pratt Sensitivity Function. *Mathematics of Operations Research* (to appear), 2015.
- [8] Chow, Y.S., Robbins, H. and Siegmund, D. *Great Expectations: The Theory of optimal Stopping*. houghton Mifflin, Boston, 1971.
- [9] Dana, R.A.: A representation result for concave Schur concave functions. *Mathematical Finance*, **15** (2005) 613-634.
- [10] Davis, M.A.H. and Lleo, S. *Risk-Sensitive Investment Management*. World Scientific, 2014.
- [11] Di Masi, G., and Stettner, L.: Risk sensitive control of discrete time partially observed Markov processes with infinite horizon. *Stochastics* **67**(3-4) (1999) 309-322.
- [12] Hardy, G.H., Littlewood, J.E. and Pólya, G. *Inequalities*. Cambridge University Press, Cambridge, 1934.
- [13] Howard R.A. and Matheson J.E.: Risk-sensitive Markov Decision Processes. *Management Science*, **18** (1972) 356-369.
- [14] James M.R., Baras, J.S. and Elliott, R.J.: Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. *IEEE Transactions on Automatic Control*, **39** (1994) 780-792.
- [15] Kadota, Y., Kurano M. and Yasuda M.: Utility-optimal stopping in a denumerable Markov chain. *Bulletin of Informatics and Cybernetics*, **28** (1996) 15-21.
- [16] Müller, A.: Certainty equivalents as risk measures. *Brazilian Journal of Probability and Statistics*, **21** (2007) 1-12.
- [17] Müller, A.: Expected utility maximization of optimal stopping problems. *European Journal of Operational Research*, **122** (2000) 101-114.
- [18] Shiryaev, A.N. *Optimal Stopping Rules*. Springer-Verlag, Berlin Heidelberg, 2008.
- [19] Stettner, L.: Risk sensitive portfolio optimization with completely and partially observed factors. *IEEE Transactions on Automatic Control*, **49** (2004) 457-464.
- [20] White, D.J.: Utility, probabilistic constraints, mean and variance of discounted rewards in Markov processes. *OR Spektrum*, **9** (1987) 13-22.
- [21] Whittle, P. *Risk Sensitive Optimal Control*. Wiley, Chichester, 1990.
- [22] Tamaki, M.: Optimal selection from a gamma distribution with unknown parameter. *Z. Oper. Res. Ser. A-B*, **28** (1984) 47-57.

(N. Bäuerle) DEPARTMENT OF MATHEMATICS, KARLSRUHE INSTITUTE OF TECHNOLOGY, D-76128 KARLSRUHE, GERMANY

*E-mail address:* nicole.baeuerle@kit.edu

(U. Rieder) UNIVERSITY OF ULM, D-89069 ULM, GERMANY

*E-mail address:* ulrich.rieder@uni-ulm.de